REPORT to ICOT

by G. Gazdar

INTRODUCTION

I was invited to visit Tokyo in December 1983 by both ICOT and IPA independently. Once they realized that they had both invited me they coordinated my visit between them. Due to inescapable pressures on my time in England, my stay in Japan had to be shorter than either ICOT or I would have preferred. During my 11 day visit, I participated in a workshop on nontransformational grammar at ICOT, gave a series of six public lectures on behalf of IPA, gave a talk at Sophia University, and had numerous discussions with Japanese linguists, computational linguists, and computer scientists from ICOT and outside ICOT. These activities are described in more detail below.

WORKSHOP AT ICOT: 12-13 December

The first activity of my visit was to participate in a workshop on non-transform ational approaches to grammar and computational linguistics at ICOT. My own paper reported recent joint research done by myself and others at UCLA this last summer on some revisions to the GPSG theory of features. It was slightly unfortunate that this talk, reporting very recent work, had to be presented before my IPA lectures were able to sketch in the relevant background for those members of the audience who were less familiar with the technical niceties of the GPSG feature system.

The quality of the papers given at the workshop was at least of the standard to be expected at a comparable workshop in the USA, and rather better than one might expect in, say, the UK. It was clear that Japanese computational linguists were very much in touch with the latest developments in syntax (e.g. GPSG and LFG) and semantics (e.g. situation semantics). The Montague tradition has taken root in Japanese computational linguistics in a way which it has not it in British computational linguistics, and I believe that this will prove very much to Japan's advantage in the long term. And it is clear that Japanese computer scientists have decided, quite rightly in my view, that the transformational tradition has little, if anything, to offer computational linguistics in terms of things that can be usefully implemented.

LECTURES FOR IPA: 14-16 December

I gave a series of six lectures, each of approximately three hours duration (including questions and discussion). These lectures amounted to a fairly detailed "crash course" in the GPSG framework as it has so far been developed. Between 40 and 50 people attended, and I gather that these were drawn about equally from the linguistics and computer science communities. Many of the attendees were graduate students, but there was, in addition, a significant group of faculty some of whom were already known to me from their work (e.g. Professors Gunji, Kajita, Muraki, and Nishida). I was naturally very pleased that such scholars chose to attend and exhibited such evident interest. What struck me very forcibly about the lectures was the excellence of the

questions and comments from the floor, questions and comments which revealed a deep understanding of what I was talking about, and a full appreciation of the issues and problems at stake. The standard of this discussion was much higher than I would expect from a comparable audience in England or the US, and thankfully free from the aggression so often evident in the anglo-american academic environment.

Mr. Murata, my IPA host, very thoughtfully arranged lunch on each day for different groups of Japanese linguists and computer scientists attending my lectures, and this provided me with a welcome opportunity to discuss things less formally with a good cross-section of my audience.

LECTURE AT SOPHIA UNIVERSITY: 17 December

At the invitation of Professors Ikeya and Ota, I presented a two hour lecture (a somewhat simplified and abbreviated version of my sixth IPA lecture) to a largely transformationalist audience at Sophia University. The lecture dealt with the GPSG analysis of some recently rediscovered facts that have provoked a spate of discussion in the recent transformationalist literature. The talk appeared to be well received, and there were some questions of the standard I had come to expect (though I noticed that these questions came largely from people who had been attending the IPA lectures).

DISCUSSION AND DEMONSTRATIONS AT ICOT: 19-21 December

The NKG group presented their work to Professor Gunji and myself, both as a group and individually. I was shown demonstrations of a number of systems including their LFG package, morphological analyzer, POPS parsing, "three wise men" (concurrent prolog demo), a grammar inference demo, and Professor Tanaka's English-to-Japanese machine translation system. I was impressed by all of these demonstration systems and learnt quite a bit from them. One consequence of these discussions and demonstrations as far as my own research is concerned, is that I now plan to use Japanese as one language in the sample of about five languages that I shall be looking at in the research on lexical representation and inflectional morphology that I shall be beginning in the late summer of 1984 at the Centre for Advanced Study in the Behavioral Sciences in Palo Alto. This will enable me to make good practical use of the expertise that I have been exposed to here.

A major benefit of my visit to ICOT has been the opportunity it has provided me of sustained conversations with Professor Gunji, a linguist that I rate extremely highly, and someone whom I have been in correspondence with for a number of years. Meeting him in person served to further enhance my respect for him. ICOT is both sensible and fortunate to have him as a member of their Working Group 3.

CONCLUDING COMMENTS

One cannot help but be impressed by the energy and enthusiasm that permeates ICOT. The four ICOT staff that I came to know best and had most to do with, namely the members of the NKG, are clearly people of the highest calibre. I am not a computer scientist, but I have no difficulty recognizing intelligent, hardworking, strongly motivated people from sister disciplines when I encounter them. The work they are engaged in seems to me to be soundly based, intellectually

interesting, and entirely consonant with overall aims of the Fifth Generation Project. I have no doubt that they will achieve most of the subgoals that they have been set for the initial three year period of the project. If they do not achieve some of these goals then that will only be due to lack of manpower and time, not lack of talent or effort.

I was surprised by the small size of the group - a very comparable range of work is done by a natural language and knowledge representation group of more than twice the size at the Palo Alto Computer Science Laboratory of the Hewlett Packard Company. The syntax subgroup of this latter itself consists of four full-time people (two linguists, two computer scientists) plus two external consultants (both linguists) who work one day a week each.

I was also surprised that the group had gone as far as it had without deciding on a grammar representation language. I think this decision is a rather fundamental one, affecting, as it does, the nature of lexical entries (and hence the dictionary project), the way semantics is most naturally construed (and hence semantic analysis and the mode of knowledge representation), and even the space of appropriate parsers. I would advise them to take a decision on this matter as soon as possible. In choosing, designing, or modifying a grammar representation language, a criterion to bear in mind is the present availability (or nonavailability) of Japanese and English grammar fragments in that representation language.

I think the group, consisting as it does of people trained primarily in computer science (at least, that's my understanding), underestimates both the difficulty and time involved in grammar construction once the grammar representation language issue is decided. My own view is that this is work that should be "farmed out" to linguistics graduate students on a piecework basis. Mr. Murata tells me that IPA standardly employs graduate students in this kind of way. Once such students have been trained in the intuitive interpretation of whatever grammar description language is chosen, they should be able to do the rather tedious grammar (and lexicon) construction work faster and more cheaply than people with a computer science degree. A fully explicit grammar for a large fragment of English or Japanese is going to be a big enterprise involving a lot of person-hours. I think it is a waste of resources to use the time of ICOT staff to do this work themselves.

Grammar construction can be speeded up by computational grammar development tools (for example, to check that a given rule actually generates the sentences that one expects it to), and I was pleased to see that some thought was being given to this matter. Once the grammar representation language is decided, then the construction (or adaptation) of such tools should be made a matter of some priority.

In the longer term, whilst it may be economic now to hire a flock of graduate students to construct grammars for English and Japanese, it clearly won't be sensible to have to do the same thing in the future for Chinese, Spanish, Russian, German, Swahili, French, and so on. So I was also glad to see that the group had been giving some thought to the issue of inferring grammars from corpora. This is bound to become an important issue within five to ten years.

One final postscript of a point (one which it probably isn't my place

to make): I was surprised that all the ICOT staff currently have to timeshare a single DEC20. During some of the demonstrations the response times were really rather slow (compared with what my AI colleagues at Sussex will now tolerate in their research). It seems to me something of a false economy to have such skilled people sitting round waiting for a machine to react. I guess the problem is only temporary and will disappear once the in-house designed personal machines become available to the staff. But I couldn't help noticing that the computing facilities at IPA were a great deal more powerful than those at ICOT.

I cannot conclude this document without taking the opportunity to express my very sincere thanks to all the staff of ICOT and IPA that I met, and to Professors Gunji and Ikeya, for making me so welcome and for making my visit such an agreeable one. Special thanks are due to Mr. Mukai (ICOT) and Mr. Murata (IPA) who extended hospitality and friendship to me that went well beyond the duties of perfect hosts.